NEWS OR NOISE?
HOW NEWS DRIVES COMMODITY PRICES

Completed Research Paper

Stefan Feuerriegel  
University of Freiburg  
Platz der alten Synagoge  
79098 Freiburg, Germany  
stefan.feuerriegel@is.uni-freiburg.de

Dirk Neumann  
University of Freiburg  
Platz der Alten Synagoge  
79098 Freiburg, Germany  
dirk.neumann@is.uni-freiburg.de

Abstract

The efficiency of electronic markets is affected by the availability of information. While some information comprises quantitative data, the majority is of qualitative nature. Although qualitative facts are essential, they are difficult to decode. As a status quo, knowledge on information processing of human agents facing qualitative news is still marginal and mostly unknown. Accordingly, it is crucial to understand how different decision makers process qualitative information in commodity markets. In this paper, we show that sentiment analysis facilitates research in qualitative information processing. We provide empirical evidence that abnormal returns in commodity markets can be explained, up to a large extent, by news sentiment. Further, we contribute to the question of how investors process novel information. Our results suggest that the influence of sentiment is asymmetric – with commodity markets being mostly driven by negative news. In addition, we show that words exhibit strong changes in polarity over time.

Keywords: Information processing, sentiment analysis, text mining
Introduction

Being fundamental parts of today’s economies, information technology has fostered the emergence of electronic markets. As a result, electronic markets have not only transformed e-commerce and stock markets, but also commodity exchanges. Prominent examples comprise the Intercontinental Exchange (ICE) and the CME Globex trading system as used by the Chicago Mercantile Exchange (CME) as well as the New York Mercantile Exchange (NYMEX) – each of them represent clearing electronic transactions with daily volumes exceeding several billion dollars. Having this relevance in mind, it is essential to understand how electronic markets perform well in their function of efficient resource allocations.

Up to large extent, market efficiency relies upon the availability of information (Fama 1965). Access to market information is promoted with ease in electronic markets and, because of the straightforward access, decision makers (i.e. consumers, suppliers and intermediaries) can use more information to make purchases and sales more beneficial (e.g. Granados et al. 2010). However, the bulk of available information has – ever since the advent of the Information Age – literally exploded. As knowledge on information processing of human agents facing (qualitative) information (e.g. Liebmann et al. 2012) is marginal and mostly unknown, we aim at closing this gap. In fact, it is both native and crucial to research how decision makers process and act upon (qualitative) information in commodity markets.

In this paper, we regard information as the combination of all relevant news that impacts prices – no matter whether of quantitative or qualitative nature. Besides quantitative information, information in the form of textual news obviously also provides valuable insights (Henry 2008; Henry and Leone 2009; Loughran and McDonald 2011; Siering and Muntermann 2013; Tetlock 2007). Hence, information processing has been subject of many IS research publications. For example, Siering (2013) examine the relationship between media sentiment and investor attention; finding that the positive impact of media sentiment on returns is increased when investor attention is high. On top of that, news sentiment can be utilized to analyze how information is processed (Siering 2013) as investors and analysts perceive novel information differently. Liebmann et al. (2012) comes to the conclusion that investors rapidly translate novel information into transactions whereas analysts wait to respond. Thus, these research questions are placed at the intersection between Information Systems and Behavioral Finance. Research on Behavioral Finance in the area of textual sentiment has been receiving great traction lately. As one of the most prominent examples, Shleifer and Summers (1990) use sentiment as a proxy for speculation as done in their so-called noise trader approach.

In this research paper, we address information processing in commodity markets. More specifically, Lechthaler and Leinert (2012) asks “what is driving the oil price”? When analyzing the driving forces in commodity markets, several external impact factors have been identified. These factors cover a broad variety ranging from economic variables such as GDP to exchange rates. While many publications from related literature determine different fundamentals that drive prices in commodity markets, there is a huge portion of residual noise that is still unexplained. Hence, it is an intriguing question to find out, to what extent, the residual noise can be explained by the textual content of news. Consequently, we
investigate how investors react to related news announcements in order to empirically study information processing in commodity markets. To extract knowledge from the textual representation of news, a common approach is to transform a news announcement into a sentiment score that indicates the polarity of its content. We use information with techniques from Information Systems (IS) research to investigate our research question. To succeed in this goal, we exploit sentiment analysis to measure the relation between investor behavior and the content in the announcements. We will focus on two different sample commodities, namely, oil and gold. The reason for the selection is that prices of both commodities (see Figure 1) show a completely different curvature and both are commodities with high demand. Albeit the commodities differ in their characteristics, the effect of news on the commodity prices is remarkably robust.

As a main contribution to IS research, this paper sheds some light into information processing in commodity markets. As prices in commodity markets are affected by other market returns, we use the so-called abnormal returns (the actual return minus the expected return) to measure the development of commodity prices. Individual research questions as well as contributions are as follows.

Research Question 1: (a) To what extent are abnormal returns of oil driven by news sentiment? (b) To what extent are abnormal returns of gold driven by news sentiment?

As a first question, we analyze the impact of news sentiment as well as a broad range of fundamental variables on crude oil prices. As a contribution to IS research, we find that most control variables show only little significance. However, abnormal returns are to a large extent determined by news sentiment. To show this finding, we review approaches for sentiment analysis and, then, adapted methods suitable for analyzing all daily announcements are contributed.

Research Question 2: To what extent are abnormal returns in commodity markets driven by news volume?

Related literature have propose news volume as a metric to identify extreme changes in oil prices. We count the number of news announcements as well as the number of words as an additional independent variable. When comparing news volume with sentiment approaches, we provide empirical evidence that sentiment approaches are more accurate in explaining movements in commodity markets.

Research Question 3: What is the influence of positive and negative news sentiment on commodity prices?

Further, we analyze the impact of positive and negative news separately. Consistent with research on capital markets, we prove that commodity markets are also driven by negative sentiment.

The remainder of this paper is structured as follows. In Section 2, we review related IS publications on information processing in commodity markets and compare approaches for sentiment analysis. Based on sentiment measures from literature, we adapt these approaches to gauge the sentiment of commodity news – giving our research model in Section 3. Finally in Section 4, we carry out an event study that measures the impact of news empirically. Section 5 concludes the paper with a summary and an outlook on future research.

Related Work

In this section, we present related literature grouped into two categories. First, studies on information processing in commodity markets are revised. Second, we compare approaches that measure news sentiment. Altogether, all following references provide evidence that studying how news drives commodity prices is a relevant and important research question to the IS community.

Information Processing in Commodity Markets

While information processing has been extensively studied in capital markets, literature focusing on commodity markets is rare. Wex et al. (2013) concludes that “even though information system methodology, i.e. text mining, is common when it comes to financial (stock) market predictions based on analyses of financial (ad-hoc) messages, the literature is remarkably silent about the oil domain.” While searching for related literature, we came across only the following few references:
Several references perform event studies to analyze the impact of OPEC announcements on oil prices. For example, Demirer and Kutan (2010) examines the informational efficiency of crude oil spot and futures markets with respect to these announcements. Lin and Tamvakis (2010) provide evidence on the effects of OPEC announcements on major international crudes by examining announcements from both official conferences and ministerial meetings. The OPEC conferences indicate and announce changes in fundamental variables (such as quotas), but this is opposed to analyzing the content of news and actual news sentiment.

Wex et al. (2012) and Wex et al. (2013) create an early warning system which aims at identifying oil-critical events. The authors count Reuters news announcements assuming that unexpected changes in the news volume indicate oil-critical events. However, the authors rely on news volume only and do not dive into the content of messages. Thus, we expect sentiment analysis to improve results and empirically validate this conjecture later.

Closest to our research question are the following two text mining approaches. Yu et al. (2005) aim at predicting the direction of oil prices changes only. The authors achieve this by using text mining to extract relevant features which are integrated in a pattern-based system. Rad (2009) uses agency news to train a Support Vector Regression that predicts the magnitude of price changes, but the author does not evaluate the impact of news. These approaches are impractical when it comes to monitoring the effect of news sentiment on the magnitude of commodity price movements.

Lechthaler and Leinert (2012) aim at analyzing the long-term effects of news sentiment along with various variables on crude oil prices. Hence, the authors construct a structural vector autoregressive model and use a sentiment time series for the crude oil obtained from the proprietary Thomson Reuters News Analytics database. However, their results are not suited for answering our research question as the authors do not provide results on the coefficients within their model; however, these are necessary to evaluate news impact. Instead of monthly resolution that measures long-term effects, we use daily data to assess news impact.

Feuerriegel et al. (2014) attempt to predict when oil bubbles may bust using the sentiment of news announcements. Accordingly, the authors first try to understand how news reception evolves depending on the market phase (boom or bust). Then, the probability of oil bubble bursts are calculated on the basis of a Markov-regime switching model.

Since the above research papers concentrate on a partial examination of news sentiment in commodity markets, many questions are left unanswered. All listed publication lack (1) a comparison of different sentiment metrics to gain robust results, and (2) an in-depth evaluation of news sentiment on stock market returns using statistical methods. Consequently, an empirical study analyzing how news sentiment impacts returns of commodities still seems to be an open research question. As a consequence, we pursue a rigorous IS approach: First, we perform an event study to compute abnormal returns. Second, we specify a model to measure the impact of news sentiment empirically.

Methods for Sentiment Analysis

Methods that use the textual representation of documents to measure the positivity and negativity of the content are referred to as opinion mining or sentiment analysis. In fact, sentiment analysis can be utilized to extract subjective information from text sources as well as to measure how market participants perceive and react upon news. Here, one uses the observed stock price reactions following the news announcement to validate the accuracy of the sentiment analysis routines. Based upon sentiment measures, one can identify the relationship between news and stock market reaction. In the finance discipline, Antweiler and Frank (2004) and Tetlock (2007) have demonstrated among the first that a discernible relation between news content and their stock market reactions exists. Within the IS community, sentiment analysis has gained significant influence and a huge number of references from various applications – for example, Groth and Muntermann (2008), Hagenau et al. (2012b), Liebmann et al. (2012), Mittermayer (2004), Oh and Sheng (2011), (Siering 2012), Tsai et al. (2010) and Yu et al. (2005) – has confirmed that sentiment analysis is both powerful and effective.

As sentiment analysis is applied to a broad variety of domains and text sources, researchers have devised various approaches to measure sentiment. Pang and Lee (2008) provide a comprehensive domain-independent survey of approaches in sentiment analysis. Within finance, recent literature surveys
(Mittermayer and Knolmayer 2006b; Minev et al. 2012) compare studies aiming at stock markets prediction. All approaches can be roughly grouped in four categories: Rule-based methods rely on small set of rules that, mostly, aim at short text sources (e.g. from social media), but are rarely used when it comes to large sources. Dictionary-based approaches (e.g. Demers and Vega 2010; Henry 2008; Jegadeesh and Di Wu 2011; Tetlock et al. 2008) are very frequently used in recent financial text mining research. These methods count the frequency of pre-defined positive and negative words – producing results that are straightforward and reliable. Term weighting approaches (e.g. Mittermayer and Knolmayer 2006a; Liebmann et al. 2012) use word frequencies from a training set (i.e. the prior probabilities) to assign a weight to each term. Subsequently, the out-of-sample set can be processed and, then, the sentiment of each news announcement is computed by aggregating term weights. Machine learning approaches (e.g. Antweiler and Frank 2004; Li 2010; Mittermayer and Knolmayer 2006a; Schumaker and Chen 2009) offer a broad range of methods, but might suffer from overfitting (Sharma and Dey 2012). In our research framework, we have to process only few data points linked with a large text basis (i.e. hundreds of announcements from a single day) along with a continuous return. Thus, we have experienced difficulties to gain robust results and focused on the above approaches instead.

All of the above approaches were originally designed to extract the sentiment of an individual news announcement. As an additional contribution, we extend all approaches to handle the news stream of a whole day by aggregating the sentiment values of all announcements. In summary, we have a broad set of techniques to measure sentiment at hand. By using several of these approaches, we will ensure the robustness of our results.

**Research Methodology: Event Study Design**

This section introduces our research methodology as depicted in Figure 2. In a first step, only those news announcements are filtered that fit our research focus. Then, each announcement is subject to preprocessing which transforms the running text into machine-readable tokens and replaces synonyms. The final step before evaluation is sentiment analysis, where all daily announcements are aggregated to compute the corresponding sentiment measure. Ultimately, we perform an event study to analyze the influence of sentiment on abnormal returns.

**Event Study Methodologies and Abnormal Returns**

Event studies use data from financial markets to inspect changes in financial values due to a specific event and measures its impact. A recent literature review (Konchitchki and O’Leary 2011) reveals how Information Systems research has introduced and exploited event study methodology – turning event study into both an effective and widespread approach. Its general procedure (MacKinlay 1997) involves the following steps: (1) identifying the events of interest; (2) defining the so-called event window; (3) predicting a normal return during the event window in the absence of the event; (4) estimating the difference between actual and normal return defined as the abnormal return.

The initial task is to define the event of interest. In our research, the event consists of the daily commodity-related announcements from the news corpus. In a second step, we identify the time interval
during which the commodity price involved in this event is examined. This period of interest is defined as the event window. Our event window, as we are provided with daily financial market data, includes the single day of the announcement stream. To extract the impact of daily events, event study defines the abnormal return. The abnormal return equals the actual return (during the event window) minus the expected return based on the pre-event time. This resulting value, i.e. the abnormal return, gives credit to the extracted effect of the events and measures its impact. In contrast to that, the normal return is defined as the expected return without conditioning on the event \( X_t \) at time \( t \) taking place. Then, the abnormal return of an arbitrary commodity \( c \) is given by

\[
AR_c(\tau) = R_c(\tau) - E(\tau|X_t) - \alpha_c - \beta_c R_m(\tau) - \epsilon_{c,t} \tag{1}
\]

where \( AR_c(\tau), R_c(\tau) \) and \( E(\tau|X_t) \) are the abnormal, actual and normal returns. The normal return is estimated during a time interval named estimation window.

According to MacKinlay (1997), we calculate the normal return by the so-called market model. This model relies on the assumptions that asset returns are jointly multivariate normal and independently and identically distributed through time. This assumption might seem strong, but is, in practice, empirically reasonable (MacKinlay 1997). The market model assumes a stable linear relation between the market return \( R_m(t) \) and the normal return, i.e. the return of the market portfolio. More precisely, the market model is for any commodity \( c \)

\[
R_c(t) = \alpha_c + \beta_c R_m(t) + \epsilon_{c,t}, \quad E(\epsilon_{c,t}) = 0, \quad \text{Var}(\epsilon_{c,t}) = \sigma^2_{c,t} \tag{2}
\]

where \( R_c(t) \) and \( R_m(t) \) are returns in period \( t \) of commodity \( c \) and on the market portfolio, respectively, and \( \epsilon_{c,t} \) is the zero mean disturbance term. Here, \( \alpha_c, \beta_c \) and \( \sigma^2_{c,t} \) are the parameters of the market model. These are determined from a regression such as ordinary least squares (OLS) based on the values from the estimation window. The higher the \( R^2 \), the greater is the variance reduction of the abnormal return – i.e. detecting the effect better. Ultimately, the abnormal return in period \( \tau \) is computed by

\[
AR_c(\tau) := R_c(\tau) - \alpha_c - \beta_c R_m(\tau). \tag{3}
\]

**Preprocessing News Announcements**

Before performing the actual sentiment analysis, several operations are involved during a preprocessing phase. The steps are as follows.

- **Tokenization.** Each announcement is split into sentences and single words named tokens (Grefenstette and Tapanainen 1994).
- **Negations.** We need to consider negations and invert the meaning of words and sentences accordingly. When encountering the word no, the weights of the subsequent three words (i.e. the object) are negated. When encountering other negating terms (rather, hardly, couldn’t, wasn’t, didn’t, wouldn’t, wouldn’t, weren’t, don’t, doesn’t, haven’t, hasn’t, won’t, hadn’t, never), all succeeding words in the sentence are. Though this approach (Dadvar et al. 2011) is simple, it is also suitable to cope with nested negations.
- **Stop words removal.** Words without a deeper meaning such as the, is, of, etc. are named stop words (Manning and Schütze 1999) and, thus, can be removed. We use a list of 571 stop words proposed in Lewis et al. (2004).
- **Synonym merging.** Synonyms, though are spelled differently, convey the same meaning. In order to group synonyms by their meaning, we follow a method that is referred to as pseudoword generation (Manning and Schütze 1999). Approximately 150 frequent synonyms or phrases from the finance domain are aggregated according to their meanings.
- **Stemming.** In computational linguistics, stemming refers to the process that reduces inflected words to their stem (Manning and Schütze 1999). One usually aims at mapping related words to the same stem, even if this stem is not itself a valid root form, as long as inflected forms are grouped together. In our analysis, we use the so-called Porter stemming algorithm (Porter 1980).
Methods for Analyzing News Sentiment

In this section, we briefly recapitulate sentiment measures for dictionary-based approaches (e.g. Tetlock $S_{\text{Tet}}(t)$ and Net-Optimism $S_{\text{NO}}(t)$) as well as term weighting metrics (e.g. Tonality $S_{\text{Ton}}(t)$ and Bi-Normal Separation $S_{\text{BNS}}(t)$). These measures were originally developed to analyze the sentiment $S(A)$ of a single news announcement. In the following sections, we present extended versions that aggregate the sentiment of all announcements from one day to result in a sentiment value $S(t)$ that represents the daily news stream. Then, let $W_{\text{tot}}(A)$ denote the total number of words in the announcement $A$; $W_{\text{neg}}(A)$ denote the number of negative words in the announcement $A$; and $W_{\text{pos}}(A)$ denote the total number of positive words in the announcement $A$.

Tetlock-Negative

Tetlock et al. (2008) define a measure for investor sentiment based on the negative wordlist of the Harvard-IV General Inquirer dictionary. The so-called Tetlock score is defined as

$$S_{\text{Tet}}(t) = -\frac{Tet(t) - \mu_{\text{Tet}}}{\sigma_{\text{Tet}}}$$

with $Tet(t) = \frac{\sum A W_{\text{neg}}(A)}{\sum A W_{\text{tot}}(A)}$, \(\mu_{\text{Tet}}\) is the mean of $Tet(t)$ and $\sigma_{\text{Tet}}$ is the standard deviation of $Tet(t)$ over a previous training subset. In fact, $Tet(t)$ measures the ratio of negative words out of all words and this score is afterwards normalized. We determine both $\mu_{\text{Tet}}$ and $\sigma_{\text{Tet}}$ once from a training interval and keep it fix explicitly such that we can measure these changes. Then, the variable $S_{\text{Tet}}(t)$ denotes the stationary measure of sentiment.

Net-Optimism

Demers and Vega (2010) and Henry (2008) propose a different score that, essentially, measures the difference between positive and negative word counts. This so-called Net-Optimism is given by

$$S_{\text{NO}}(t) = \frac{\sum A W_{\text{pos}}(A) - W_{\text{neg}}(A)}{\sum A W_{\text{tot}}(A)} \in [-1, +1].$$

Thus, Net-Optimism measures the difference between the count of positive and negative words normalized by the number of total words.

Tonality

The Tonality approach (Liebmann et al. 2012) is based on the differences between the observed relative frequency and the expected conditional probability. The weight of an individual word in a training set is given by

$$Ton(w) = \frac{1}{2} \left[ \frac{O_{\text{pos}}(w) - E_{\text{pos}}(w)}{E_{\text{pos}}(w)} - \frac{O_{\text{neg}}(w) - E_{\text{neg}}(w)}{E_{\text{neg}}(w)} \right] \in [-1, +1].$$

Here, $O_{\text{pos}}(w)$ and $O_{\text{neg}}(w)$ denote observed frequency of positive/negative announcements out of all announcements containing the word $w$. Both $E_{\text{pos}}(w)$ and $E_{\text{neg}}(w)$ denote the expected conditional probability that a word occurs in a positive/negative message based on its occurrences in the overall corpus. Tonality is greater than zero for a positive word and below zero for a negative one. Thus, the Tonality score in the out-of-sample set is

$$S_{\text{Ton}}(t) = \sum A \frac{1}{W_{\text{rel}}(A)} \sum_{w \in A} Ton(w) \cdot f(w, A) \in [-1, +1]$$

and $W_{\text{rel}}(A) = \{w \in A | \delta < |Ton(w)|\}$.

Bi-Normal Separation (BNS)

Based on previous work (Forman 2002, 2003, 2008), we propose a new term weighting approach using the Bi-Normal Separation metric. While original work employs Bi-Normal Separation for feature selection.
in machine learning, we found robust and reliable results when using the formula for term weighting. In a first step, each word is assigned a weight using a Bi-Normal Separation formula. Let $F^{-1}(x)$ denote the inverse cumulative probability function of standard normal distribution. Then, the Bi-Normal Separation weight of an individual word $w$ is defined as

$$BNS(w) = F^{-1}\left(\frac{O_{pos}(w)}{W_{pos}}\right) - F^{-1}\left(\frac{O_{neg}(w)}{W_{neg}}\right) \in ]-\infty, +\infty[. \tag{8}$$

To avoid the undefined value $F^{-1}(0)$, Forman (2002) substitutes zero by a very small number. In the second stage, we compute the sentiment score via

$$S_{BNS}(t) = \sum_{A} \frac{1}{W_{tot}(A)} \sum_{w \in A} BNS(w) \in ]-\infty, +\infty[. \tag{9}$$

Forman (2002) describes the original BNS formula as: “For intuition, suppose the occurrence of a given feature in each document is modeled by the event of a random normal variable exceeding a hypothetical threshold. The prevalence rate of the feature corresponds to the area under the curve past the threshold. If the feature is more prevalent in the positive class, then its threshold is further from the tail of the distribution than that of the negative class. The BNS metric measures the separation between these thresholds.”

**Empirical Evaluation: Measuring the Impact of News**

Having discussed the elements of sentiment analysis, we apply the analysis to measure the impact of news. Thus, we specify the news corpus and describe the regression design that inspects information processing in commodity markets. The acquired results are used for evaluating the above research questions.

**News Corpus**

Our news corpus originates from the Thomson Reuters News Archive for Machine Readable News. We choose Reuters news deliberately due to the following four reasons: (1) Reuters conveys, in particular, news about commodity markets and provides a very detailed coverage of related events. (2) Reuters news is third-party content and, thus, do not originate from market participants themselves. (3) Announcements from Reuters solely contain novel information. In comparison to announcements, articles from newspapers might be edited, perturbed, shortened and filtered by subjective criteria of editors. (4) Articles from newspapers are normally built upon announcements from news agencies. Thus, news released from newspapers might be delayed and show a time lag. Overall, choosing Reuters news provides the advantage of a more objective news source.

All announcements provided by Reuters arise from the time span January 1, 2003 till May 31, 2012. The announcements come along with additional labels indicating their content. Based upon these labels, the news corpus is filtered such that we extract announcement focusing on both oil and gold markets respectively. This is achieved by applying a set of filter criteria: (1) The language must be English. (2) The event type is Story Take Overwrite to guarantee that we not yield an alert but the actual message. (3) Special types of announcements such as alerts or personal opinions might have limited relevance to changes in commodity markets and we want to exclude these. Thus, we omit announcements that contain specific words (advisory, chronology, corrected, feature, diary, instant view, analysts view, newsmaker, corrected, refile, rpt, schedule, table, service, alert, wrapup, imbalance, update) in their headline. (4) We use topic codes (Reuters 2013) to select announcements that deal with a particular commodity. In detail, the label CRU filters all news covering crude oil and GOL filters gold-related news. (5) We exclude announcements addressing changes in prices to avoid simultaneity (Antonakis et al. 2010; Antonakis et al. 2014). (6) In order to remove white noise, we require announcements to count at least 50 words. All in all, this set of criteria filters a total of 339,446 announcements related to crude oil and 184,255 related to gold.

Now, we specify all parameters inside the sentiment analysis routines. We use the years 2003 and 2004 as training intervals. This period is used to compute the mean $\mu_{tet}$ of $tet(t)$ and the standard deviation $\sigma_{tet}$
of \( T(t) \). During the same period, both Bi-Normal Separation and Tonality determine their term weights. Similar to \{Liebmann 2012 #10\}, we set the Tonality threshold \( \delta \) to 0.05 and set the Tonality score of all words that occur fewer than 100 times to zero – this helps to discriminate word polarity.

We gained the results by performing both sentiment analysis and regression evaluation in an automated, parallelized procedure. Given the huge amount of news announcements stored in the Reuters news corpus, the runtime exceeded 10 hours on a 6-core machine while processing more than 200 gigabyte of data.

Regression Design

In this section, we investigate how investors react to related news announcements and analyze information processing in commodity markets empirically. To succeed in this goal, we exploit sentiment analysis to measure the relation between investor behavior and the content in the announcements.

In a first step, we need an appropriate statistical model. In literature, commodity prices are examined by a variety of models ranging from linear regression to time series approaches. Such complex variants are vector autoregression (VAR) models or vector error correction (VECM) models. The former captures the linear interdependencies among multiple time series; the latter adds a way to estimate both short-term and long-term effects of one time series on another. Instead of utilizing models such as VAR or VECM, we surpass issues from stationarity and cointegration by using a linear model with abnormal returns. Nevertheless, literature gives also evidence that linear models might be suitable as Gileva (2010) uses a linear model with oil and it seems to be a common approach when analyzing gold as proven in (e.g. Cai et al. 2001; Christie–David et al. 2000; Ghosh et al. 2004).

Instead of classical commodity prices, we perform an event study to extract the effect of individual events. Let \( AR_c(t) \) define the abnormal return of a commodity \( c \). Here, we distinguish two markets models. In case of oil, we follow related literature on oil-related event studies (Demirer and Kutan 2010; Draper 1984; Kozicki et al. 2012) and, thereby, model the market portfolio using a commodity index, namely the Dow Jones–UBS Commodity Index. The Dow Jones–UBS Commodity Index (formerly Dow Jones–AIG Commodity Index) is a highly liquid and diversified reference for the commodity market consisting of around twenty physical commodities and is used in the model as a proxy for the commodities market performance. The actual oil price comes from the benchmark oil price in the US which is the West Texas Intermediate (WTI). In case of gold, we use prices from the London gold fixing, while the market performance in the gold market is modeled via the FTSE100. We like to point out that we gain more observations when analyzing gold as the London gold fixing is determined twice a day. When computing abnormal returns, we use, as proposed by Wirl and Kujundzic (2004), an event window of 10 trading days prior to the event.

As we use a linear model, the key independent variable is the sentiment metric \( S(t) \). We also incorporate a set of control factors to check for internal effects and autocorrelations. All control variables are given in the next section. Then, we can specify the regression model with error terms \( \epsilon_t \) by

\[
AR_{c,log}(t) = \beta_0 + \beta_1 S(t) + \beta_2 \alpha_c + \beta_3 CAR + \sum_{t} \beta_{t+3} CV_{c,t}\epsilon_t + \epsilon_t.
\]  

(10)

Both the abnormal returns \( AR_{c,log}(t) \) as well as all control variables \( CV_{c,t}\epsilon_t \) consist of standardized log-returns. In addition to that, we add monthly dummy variables. Consequently, these dummies consider additional external events not covered by the control variables and, further, they handle non-seasonally-adjusted time series. Finally, we give justice to extreme stock price effects and remove outliers at the 1% level.

Control Variables

When isolating and extracting the effect of news, we use a wide range of control variables. These control variables originate from a broad portfolio of research papers that formulate pricing models for both oil and gold. As these publications rely on actual prices instead of abnormal returns, it is impossible to inject news sentiment into a single model from literature. Hence, we collected control variables from literature and, altogether with news sentiment, test for a possible influence. Individual variables arise from
economic key figures, exchange rates, stock market variables as well as oil- and gold-related values. By doing so, we want to assure that our results measure the impact that comes from news sentiment only and we avoid perturbations by other causes such as changes in fundamental variables or external events. Table 10 in Appendix A shows all control variables injected in the price models. If not indicated differently, all variables come from Datastream. All variables are grouped in several categories to test for external effects from (1) economic variables, (2) exchange rates, (3) market variables (i.e. stock market indices, commodities, real estate), and (4) oil-related as well as gold-related variables.

**How News Sentiment Drives Oil Markets**

We use the above regression design to provide empirical evidence that news sentiment influences abnormal returns of oil prices and we measure the magnitude empirically.

Research Question 1: (a) To what extent are abnormal returns of oil driven by news sentiment?

Our results are presented in Table 3 which analyzes a specific sentiment measure (namely, Net-Optimism with Henry’s Finance-Specific Dictionary) and states the regression results. The selection of Net-Optimism is justified by Table 4. Table 4 compares the impact of news sentiment across several approaches to ensure that our results are both valid and robust. As a result, we find that Net-Optimism along with Henry’s Dictionaries achieves the highest significance. We tested for autocorrelation, heteroskedasticity and normally distributed residuals at the 0.001 % level to ensure that the results are not confounded. When checking Variance Inflation Factors and the condition number of the matrix, we also see no indication of multicollinearity. Independence across announcements is given as long as all announcements are entirely novel and not based on an interrelated course of events. Under the assumption that commodity returns are jointly multivariate normal as well as independently and identically distributed through time, the model can be estimated using Ordinary Least Squares (OLS).

According to Table 3, we observe that, besides the alpha value from the market model, news sentiment influences abnormal returns significantly. When additionally comparing the coefficients of fundamental variables and news sentiment, we find that news sentiment (coefficient of 1.26) exceeds all other coefficients originating from fundamental variables strongly. Further, we notice a low $R^2$ value of 0.382. A possible explanation is given in Tetlock et al. (2008) since “very few control variables predict next-day returns”. Overall, this provides empirical evidence that abnormal returns are described by daily news up to a large extent.

Consequently, measuring news sentiment shows a possible path how so far noisy abnormal returns can be explained. To ensure robustness, we employ a broad range of sentiment approaches – see the results in Table 4. Given that all sentiment measures show the same sign and given that most come along with a high significance level, this contributes to the validity and robustness of our results.
Table 3. Oil Regression for Net-Optimism with Henry’s Finance-Specific Dictionary

<table>
<thead>
<tr>
<th>Measure</th>
<th>Obs.</th>
<th>Coeff.</th>
<th>Std. Err.</th>
<th>t-Val.</th>
<th>p-Val.</th>
<th>Adj. R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_{t+1}</td>
<td>2354</td>
<td>0.3014</td>
<td>0.0570</td>
<td>5.289</td>
<td>&lt; 0.001</td>
<td>0.1833</td>
</tr>
<tr>
<td>S_{t+1}</td>
<td>2354</td>
<td>0.2504</td>
<td>0.0503</td>
<td>4.979</td>
<td>&lt; 0.001</td>
<td>0.1820</td>
</tr>
<tr>
<td>S_{t+1}</td>
<td>2354</td>
<td>1.2566</td>
<td>0.0452</td>
<td>27.755</td>
<td>&lt; 0.001</td>
<td>0.3819</td>
</tr>
<tr>
<td>S_{t+1}</td>
<td>2354</td>
<td>0.3301</td>
<td>0.0535</td>
<td>6.172</td>
<td>&lt; 0.001</td>
<td>0.1869</td>
</tr>
<tr>
<td>S_{t+1}</td>
<td>2110</td>
<td>0.3059</td>
<td>0.0671</td>
<td>4.556</td>
<td>&lt; 0.001</td>
<td>0.1805</td>
</tr>
<tr>
<td>S_{t+1}</td>
<td>2110</td>
<td>0.2681</td>
<td>0.0708</td>
<td>3.787</td>
<td>&lt; 0.001</td>
<td>0.1796</td>
</tr>
</tbody>
</table>

How News Sentiment Drives Gold Markets

In addition to oil, we also use the above regression design to study the relationship between news sentiment and abnormal returns of gold prices as well as to estimate its magnitude.

Research Question 1: (b) To what extent are abnormal returns of gold driven by news sentiment?

Again, all results for Net-Optimism with Henry’s Dictionary as the specific sentiment measure with the highest significance are presented in Table 5. Robustness across various sentiment approaches is validated in Table 6. We tested for autocorrelation, heteroskedasticity and normally distributed residuals at the 0.01% level to ensure that results are not confounded. Again, we checked Variance Inflation Factors and the condition number of the matrix to find no indication of multicollinearity. With the

Thirty Fourth International Conference on Information Systems, Milan 2013
additional assumptions on announcements from above, the model can be estimated using Ordinary Least Squares (OLS).

According to Table 5, we observe that, besides the alpha value from the market model, news sentiment influences abnormal returns significantly. When additionally comparing the coefficients of fundamental variables and news sentiment, we find that news sentiment (coefficient of 0.35) strongly exceeds all other coefficients originating from fundamental variables. Overall, this provides empirical evidence that abnormal returns are described by daily news up to a large extent. Hence, measuring news sentiment shows a possible path how so far noisy abnormal returns can be explained. To ensure robustness, we employ a broad range of sentiment approaches – see the results in Table 6. Given that all sentiment measures show the same sign and given that most come along with a high significance level, this contributes to the validity and robustness of our results.

| Table 5. Gold Regression for Net-Optimism with Henry’s Finance-Specific Dictionary |
|---------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                                 | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       | (8)       | (9)       | (10)      | (11)      |
| Intercept 𝛼₀                  | -0.18 (-1.45) | -0.18 (-1.39) | -0.02 (-0.16) | -0.691*** (-4.13) | 0.77*** (3.66) | 24.43*** (4.89) | 0.15 (1.77) | 0.38*** (3.61) | 0.35*** (3.41) | 0.35*** (3.59) | 0.15** (2.58) |
| Sentiment S(t)                | 0.35*** (24.63) | 0.35*** (24.74) | 0.36*** (28.35) | 0.36*** (27.94) | 0.36*** (27.94) | 0.36*** (27.94) | 0.36*** (27.94) | 0.35*** (26.87) | 0.35*** (26.8) | 0.35*** (26.8) |
| Cum. abnorm. return CAR       | -0.04** (-3.26) | 0.02 (1.71) | 0.02 (1.88) | 0.02 (1.88) | 0.02 (1.88) | 0.02 (1.88) | 0.02 (1.82) | 0.00 (1.89) | 0.00 (1.89) | 0.00 (1.89) |
| αgold from Market Model       | -0.42*** (-33.4) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) | -0.41*** (-33.19) |
| U.S. Consumer Price Index     | -18.79*** (-4.11) | 1.04*** (4.28) | 18.13*** (4.89) | -0.04 (-1.05) | -0.01 (-0.27) | -0.02 (-0.43) | -0.02 (-0.45) | -0.12* (-2.45) | 0.82* (2.55) | 0.15 (0.67) |
| Real Eff. Exch. Rate Dollar   | -3.37*** (-4.13) | -16.85*** (-4.79) | 0.46 (1.90) | 0.15 (0.81) | 0.18 (0.90) | 0.20 (0.90) | 0.20 (0.90) | 0.82* (2.55) | 0.15 (0.67) | 0.18 (0.67) |
| U.S. Gold Ind. Capacity Util. | 11.3*** (4.90) | 0.36*** (3.69) | 0.29*** (3.23) | 0.28*** (3.05) | 0.29*** (3.22) | 0.29*** (3.22) | 0.29*** (3.22) | 0.29*** (3.22) | 0.29*** (3.22) | 0.29*** (3.22) |
| U.S. Unemp. Rate              | 0.36*** (4.89) | 0.26*** (3.9) | 0.25*** (3.79) | 0.26*** (3.94) | 0.26*** (3.94) | 0.26*** (3.94) | 0.26*** (3.94) | 0.26*** (3.94) | 0.26*** (3.94) | 0.26*** (3.94) |
| Real Eff. Exch. Rate Yen      | 0.10*** (2.93) | 0.10*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) | 0.11*** (3.04) |
| Silver Price                  | 0.10*** (8.63) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) | 0.10*** (8.57) |
| S&P 500                       | -0.02* (-2.10) | -0.02* (-2.10) | 0.18* (2.01) |
| Gold Mining Cap. Util.        |                       |                       |                       |                       |                       |                       |                       |                       |                       |                       |
| Control variables             | ✓          | ✓          | ✓          | ✓          | ✓          | ✓          | ✓          | ✓          | ✓          | ✓          |
| Adj. R²                       | 0.1076     | 0.1095     | 0.2826     | 0.2818     | 0.2818     | 0.2818     | 0.2818     | 0.2818     | 0.2818     | 0.2818     |

Stated: OLS coefficients, t-statistics in parenthesis
Observations: 4717
Dummyest: Monthly
Significance: *0.05 **0.01 ***0.001

### Table 6. Checking Robustness of Gold-Related News Sentiment; Each OLS Regression Uses All Above Control Variables as well as Monthly Dummy Variables

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S₁₁(t)</td>
<td>4717</td>
<td>0.0565</td>
<td>0.0129</td>
<td>4.370</td>
<td>&lt; 0.001</td>
<td>0.1880</td>
</tr>
<tr>
<td>S₁₀₁(t)</td>
<td>4717</td>
<td>0.0330</td>
<td>0.0148</td>
<td>2.231</td>
<td>0.026</td>
<td>0.1856</td>
</tr>
<tr>
<td>S₁₀₃(t)</td>
<td>4717</td>
<td>0.3475</td>
<td>0.0130</td>
<td>26.804</td>
<td>&lt; 0.001</td>
<td>0.2948</td>
</tr>
<tr>
<td>S₁₀₃₃(t)</td>
<td>4717</td>
<td>0.0374</td>
<td>0.0141</td>
<td>2.658</td>
<td>0.008</td>
<td>0.1887</td>
</tr>
<tr>
<td>S₁₃₃(t)</td>
<td>3730</td>
<td>0.1788</td>
<td>0.0131</td>
<td>13.647</td>
<td>&lt; 0.001</td>
<td>0.2188</td>
</tr>
<tr>
<td>S₁₃₃₃(t)</td>
<td>3730</td>
<td>0.2227</td>
<td>0.0133</td>
<td>16.772</td>
<td>&lt; 0.001</td>
<td>0.2322</td>
</tr>
</tbody>
</table>

### Testing for Causality

Ultimately, we need to overleap the limitations of the above regression to prove a significant relationship between news sentiment and abnormal returns and argue for causality. Although this hypothesis is frequently assumed, it is investigated rarely (cf. Loughran and McDonald 2011). For example, Wex et al. (2013) expect that “that oil price changes are mirrored in an (anomalous) release of news messages before the price fluctuation actually occurs”. Hence, we need to provide statistical evidence on causality in several ways. First, the chronology of events supports the hypothesis that returns are a consequence of news – i.e. publishing novel news as an event prior to stock market reaction as an effect. Second, we perform a Granger causality test. More specifically, logarithmic abnormal returns of oil prices Granger-causes news sentiment with a P-value of 0.11, whereas sentiment Granger-causes logarithmic abnormal returns with a P-value of $7.7 \cdot 10^{-15}$ which is much lower – i.e. it is more likely that news predicts abnormal returns than vice versa. In case of Gold, the difference in magnitude is similar. Third, we tested cointegration via the Augmented-Dickey-Fuller test on differences (P-value is much lower than 0.01) to find that logarithmic abnormal returns and news sentiments share a common stochastic drift. Overall, this gives strong evidence that correlations are structural and not spurious. Therefore, it is likely that news sentiment causes abnormal returns.

### Does News Volume Coincide with Returns in Commodity Markets?

Besides the previous sentiment measure, we test how changes in commodity prices are subject to news volume.

**Research Question 2: To what extent are abnormal returns in commodity markets driven by news volume?**

Here, we distinguish two approaches. On one hand, we follow the idea by Wex et al. (2012) as well as Wex et al. (2013) and count the number of daily messages $V_{msg}(t)$. On the other hand, we contemplate the different lengths of announcements and, hence, use the number of words $V_w(t)$ per day. Thus, we adjust our linear model from Equation (10) such that

$$AR_{c,log}(t) = \beta_0 + \beta_1 V(t) + \beta_2 \alpha + \beta_3 \text{CAR} + \sum_i \beta_{i+3} G c_i(t) + \epsilon_t.$$  \hspace{1cm} (11)

We perform the above regression with monthly dummies, log-return control variables from Table 10 and outlier removal at the 1% level to gain the results from Table 7. In brevity, abnormal returns in commodity markets coincide with news volume strongly; however, most sentiment approaches exhibit a significantly higher t-value and, hence, show a stronger influence. On top of that, arguing for causality – i.e. less news volume leads to decreasing abnormal returns – will be difficult or even infeasible.
### Table 7. OLS Regression Results for News Volume

<table>
<thead>
<tr>
<th>News Volume Measure</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>t-Value</th>
<th>p-Value</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil - Number of announcements</td>
<td>-0.1747</td>
<td>0.0605</td>
<td>-2.889</td>
<td>0.004</td>
<td>0.1795</td>
</tr>
<tr>
<td>Oil - Number of words</td>
<td>-0.1730</td>
<td>0.0607</td>
<td>-2.850</td>
<td>0.004</td>
<td>0.1779</td>
</tr>
<tr>
<td>Gold - Number of announcements</td>
<td>-0.0783</td>
<td>0.0126</td>
<td>-6.245</td>
<td>&lt;0.001</td>
<td>0.1953</td>
</tr>
<tr>
<td>Gold - Number of words</td>
<td>-0.0815</td>
<td>0.0126</td>
<td>-6.476</td>
<td>&lt;0.001</td>
<td>0.1958</td>
</tr>
</tbody>
</table>

**Good News or Bad News – What is the Driving Force?**

Empirical evidence on momentum and reversals (e.g. Hong et al. 2000) as well as implications of Behavioral Finance theories suggests that the influence of sentiment appears to be asymmetric. According to prospect theory (Kahneman and Tversky 1979), “the value function is normally concave for gains, convex for losses and generally steeper for losses than for gains”. In other words, stock market returns are mostly driven by negative news.

**Research Question 3:** What is the influence of positive and negative news sentiment on commodity prices?

Brown and Cliff (2005) argues that this effect originates from limits to arbitrage: “Practical limitations to short-selling activity may make it difficult for rational investors to prevent market prices from being pushed above their intrinsic value during periods of excessive optimism. On the other hand, when some investors are especially pessimistic, no similar frictions prevent arbitrageurs from taking the necessary long position.” Within the context of stock market announcements, Tetlock (2007) measures the correlation of negative and positive words to find that stock returns are highly dependent on negative words. To be consistent with these findings, we also evaluate the following linear model

\[
AR_{c,\log}(t) = \beta_0 + \beta_{neg}S_{neg}(t) + \beta_{pos}S_{pos}(t) + \beta_2 \alpha + \beta_3 CAR + \sum_{i} \beta_{i+3}CV_{ci}(t) + \epsilon_i
\]

(12)

to accommodate the influence measured as the ratio of negative words \(S_{neg}(t)\) and positive words \(S_{pos}(t)\) from the Henry’s Finance-Specific Dictionary out of all words separately. This is identical to comparing each part of the Net-Optimism formula individually. The regression outcomes are presented in Table 8. The results show a higher impact of negative words than of positive words. This is supported by both the actual coefficient and its significance. Thus, the results in the commodity market are consistent with Tetlock (2007) – the influence of sentiment is asymmetric.

### Table 8. OLS Regression Results for Positive and Negative Words Separately

<table>
<thead>
<tr>
<th>Measure</th>
<th>Coefficient</th>
<th>Std. Err.</th>
<th>t-Value</th>
<th>p-Value</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil - Sentiment of negative words</td>
<td>-427.2</td>
<td>20.89</td>
<td>-20.452</td>
<td>&lt;0.001</td>
<td>0.3686</td>
</tr>
<tr>
<td>Oil - Sentiment of positive words</td>
<td>223.2</td>
<td>18.04</td>
<td>12.326</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Gold - Sentiment of negative words</td>
<td>-59.4</td>
<td>2.982</td>
<td>-19.908</td>
<td>&lt;0.001</td>
<td>0.2914</td>
</tr>
<tr>
<td>Gold - Sentiment of positive words</td>
<td>34.2</td>
<td>2.452</td>
<td>13.446</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

**Methodological Challenges: Coping with Polarity Changes**

Investors can change their option on how to assess news facts which, hence, affects their perception of information and, in particular, words. For example, the word Internet might have been initially regarded as a positive word, but this changed during the dot-com bubble. In order to measure how perception alters, we propose the following approach. Here, we utilize the score assigned to each word as a metric to identify polarity changes. We employ feature extraction metrics given by Bi-Normal Separation as a
possible path to measure changes in the polarity of words. A few examples illustrate the descriptive nature as depicted in Figure 9. The word crisis shows, in the oil-context, a high volatility in late 2008 coinciding with a simultaneous price fall. In the case of gold, the polarity of FED increases strongly. This might be a result of investor switching to gold whenever the federal funds rate decreased. Obviously, the current approach for creating dynamic dictionaries is actually misleading, as they create a dictionary based on past observations. Once it is determined it stays fixed and, thus, shares the same limitations as static dictionaries. True dynamic dictionaries are consequently necessary in the future.

![Figure 9. Polarity of Selected Words from Commodity-Related News](image)

**Conclusion and Outlook**

Ever since the advent of the Information Age, the bulk of available news has boosted rapidly. News consists of a few quantitative facts and mostly qualitative information. The information processing of human agents facing qualitative (i.e. textual) news is an active research question in the context of efficient electronic markets (e.g. Liebmann et al. 2012), though knowledge is rare. To close this gap, research must combine information inequality, subjectivity of decision makers and increased processing costs of qualitative information that impede decision efficiency. Hence, it is both native and crucial to research how decision makers process and act upon qualitative information in commodity markets.

In this paper, we show that sentiment analysis successfully accompanies research in information processing of qualitative information. Hence, we derived various metrics to measure the sentiment in commodity-related news. We chose a broad range of approaches to verify the robustness of our results. As a result, we have succeeded in demonstrating the influence of commodity-related news on abnormal returns. Achieving this goal unveils a possible path of how the impact of news on prices can be measured. Our empirical evidence reveals that, although news volume correlates with abnormal returns, sentiment has a more significant impact. However, the influence of news sentiment is asymmetric. In fact, our results suggest that commodity markets are mostly driven by negative news. Although negative news shows a stronger effect than positive news, the polarity of news is not static at all, but changes over time. We proposed an approach to identify words that face strong changes in polarity and, consequently, showed how investors change their perceptions over time.

In future work, we will advance the above methods to gain a framework that can distinguish between short-term and long-term trends of news sentiment not only on abnormal return but also on actual commodity prices. To surpass these limitations of our event study, we need to use approaches from time series analysis such as vector autoregressive models (VAR) or vector error correction models (VECM). In addition, we aim at pursuing a different direction where we not only explain movements in commodity prices but use news sentiment to predict price changes. In order to predict future prices, we need further effort to integrate machine learning techniques into the above approach. An alternative idea (Mogotsi 2007) is to detect individual trends that drive prices. While each news announcement is assigned an equal weight in the above approach, a different research question focuses on assigning a weight to each announcement.
## Appendix A: Control Variables

### Table 10. Control Variables in Price Model

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Freq.</th>
<th>Details</th>
<th>Reference</th>
<th>For Oil</th>
<th>For Gold</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Economic Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US GDP</td>
<td>Quarterly</td>
<td>Proportional quarter-to-quarter change</td>
<td>Chen and Chen (2007); Pirog (2005); similar ideas by Gileva (2010); Lechthaler and Leinert (2012); Cai et al. (2001); Christie–David et al. (2000) for Gold</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>US Inflation Rate</td>
<td>Monthly</td>
<td>Source: Federal Reserve</td>
<td>Chen and Chen (2007); Chatrath et al. (2012); Gileva (2010); Miller and Ratti (2009)</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>US Consumer Price Index</td>
<td>Monthly</td>
<td>Not seasonally adjusted; all urban consumers</td>
<td>Baker and Tassel (1985); Cai et al. (2001); Christie–David et al. (2000); Fortune (1987); Ghosh et al. (2004); Koutsoyiannis (1983)</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>US Money Stock M1</td>
<td>Weekly</td>
<td>Source: Federal Reserve Economic Data; seasonally adjusted</td>
<td>Tandon and Urich (1987) for Gold</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>US Interest Rate</td>
<td>Weekly</td>
<td>3 Month US T-Bill Rate</td>
<td>Baker and Tassel (1985); Koutsoyiannis (1983); Tully and Lucey (2007)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Treasure Bond 5 Years</td>
<td>Monthly</td>
<td>5-Year Treasury Constant Maturity Rate</td>
<td></td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>US Unemployment Rate</td>
<td>Monthly</td>
<td>Seasonally adjusted</td>
<td>Chatrath et al. (2012); Chen and Chen (2007) for Oil and Cai et al. (2001); Christie–David et al. (2000) for Gold</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>US Producer Price Index</td>
<td>Monthly</td>
<td>Finished Goods; seasonally adjusted</td>
<td>Chatrath et al. (2012) for Oil and Cai et al. (2001); Christie–David et al. (2000); Tandon and Urich (1987) for Gold</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Household Debt</td>
<td>Quarterly</td>
<td>Household Credit Market Debt Outstanding; Seasonally adjusted</td>
<td>Additionally</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td><strong>Exchange Rates</strong></td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Real Effective Exchange Rate Dollar</td>
<td>Monthly</td>
<td>Index; not seasonally adjusted</td>
<td>Amano and van Norden (1995); Bencivenga et al. (2012); Chen and Chen (2007); Pirog (2005) for Oil and Baker and Tassel (1985); Koutsoyiannis (1983); Tully and Lucey (2007) for Gold</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Real Effective Exchange Rate Yen</td>
<td>Monthly</td>
<td>Index; not seasonally adjusted</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Real Effective Exchange Rate Euro</td>
<td>Monthly</td>
<td>Index; not seasonally adjusted</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Real Effective Exchange Rate Pound</td>
<td>Monthly</td>
<td>Index; not seasonally adjusted</td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Market Variables (Stock Market, Commodities, Real Estate)</strong></td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>S&amp;P 500 Index</td>
<td>Daily</td>
<td></td>
<td>Gileva (2010); Miller and Ratti (2009)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Wheat Prices</td>
<td>Daily</td>
<td>Wheat No. 2, Soft Red</td>
<td>Additionally</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>House Price Index</td>
<td>Monthly</td>
<td>Case Shiller 20 Index</td>
<td>Additionally</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Gold Price</td>
<td>Daily</td>
<td>London Gold Fixing; Afternoon</td>
<td>Bencivenga et al. (2012)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Crude Oil Price</td>
<td>Daily</td>
<td>WTI</td>
<td>Koutsoyiannis (1983)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Oil-Related Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Oil Production</td>
<td>Monthly</td>
<td>Crude Oil Production World in Barrel/Day</td>
<td>Additionally</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>US Oil Imports</td>
<td>Weekly</td>
<td>US Imports of Crude Oil; Source: Energy Information Administration</td>
<td>Bencivenga et al. (2012)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>US Oil Refinery Utilization</td>
<td>Weekly</td>
<td>Rate in Percent; Source: Energy Information</td>
<td>Dëes et al. (2008)</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Administration</td>
<td>US Oil Inventories</td>
<td>Monthly</td>
<td>Exclusive Strategic Petroleum Reserve</td>
<td>Dées et al. (2008); Pirog (2005)</td>
<td>✓</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------</td>
<td>---------</td>
<td>-------------------------------------</td>
<td>----------------------------------</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>US Oil Capacity Utilization</td>
<td>Quarterly</td>
<td></td>
<td>Chatrath et al. (2012); Dées et al. (2008); Kaufmann et al. (2004)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>Oil Future Contracts</td>
<td>Daily</td>
<td>NYMEX-WTI Crude Oil Swap; Volume Traded</td>
<td>Bencivenga et al. (2012); Wirl (2008)</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Gold-Related Variables**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US Industry Capacity Utilization</td>
<td>Monthly</td>
<td>Rate; Seasonally adjusted</td>
<td>Christie–David et al. (2000)</td>
<td>✓</td>
</tr>
<tr>
<td>Gold Mining Capacity Utilization</td>
<td>Monthly</td>
<td>Index; Seasonally adjusted</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

**References**


Reuters 2013. “Reuters Codes: A quick guide,”


